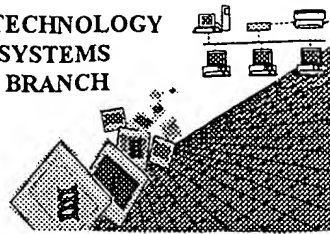


BIOTECHNOLOGY
SYSTEMS
BRANCH



RAW SEQUENCE LISTING
ERROR REPORT

#19
1652
RECEIVED
FEB 06 2002
TECH CENTER 1600/2900

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: 09/586,744A
Source: 1645
Date Processed by STIC: 1/24/2002

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.

PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:

- 1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
- 2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY

FOR CRF SUBMISSION QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.

FOR SEQUENCE RULES INTERPRETATION, PLEASE CONTACT ROBERT WAX, 703-308-4216.

PATENTIN 2.1 e-mail help: patin21help@uspto.gov or phone 703-306-4119 (R. Wax)

PATENTIN 3.0 e-mail help: patin3help@uspto.gov or phone 703-306-4119 (R. Wax)

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE **CHECKER** **VERSION 3.1 PROGRAM**, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW FOR ADDRESS:

<http://www.uspto.gov/web/offices/pac/checker>

Applicants submitting genetic sequence information electronically on diskette or CD-Rom should be aware that there is a possibility that the disk/CD-Rom may have been affected by treatment given to all incoming mail.

Please consider using alternate methods of submission for the disk/CD-Rom or replacement disk/CD-Rom.

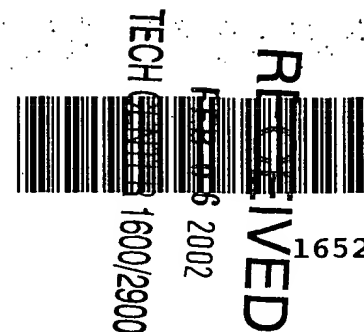
Any reply including a sequence listing in electronic form should NOT be sent to the 20231 zip code address for the United States Patent and Trademark Office, and instead should be sent via the following to the indicated addresses:

1. EFS-Bio (<<http://www.uspto.gov/ebc/efs/downloads/documents.htm>> , EFS Submission User Manual - ePAVE)
2. U.S. Postal Service: U.S. Patent and Trademark Office, Box Sequence, P.O. Box 2327, Arlington, VA 22202
3. Hand Carry directly to:
U.S. Patent and Trademark Office, Technology Center 1600, Reception Area, 7th Floor, Examiner Name, Sequence Information, Crystal Mall One, 1911 South Clark Street, Arlington, VA 22202
Or
U.S. Patent and Trademark Office, Box Sequence, Customer Window, Lobby, Room 1B03, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202
4. Federal Express, United Parcel Service, or other delivery service to: U.S. Patent and Trademark Office, Box Sequence, Room 1B03-Mailroom, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202

Revised 01/29/2002

Raw Sequence Listing Error Summary

ERROR DETECTED	SUGGESTED CORRECTION	SERIAL NUMBER: 09/586,744A
ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE		
1 _____ Wrapped Nucleics Wrapped Aminos	The number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor after creating it. Please adjust your right margin to .3; this will prevent "wrapping."	
2 _____ Invalid Line Length	The rules require that a line not exceed 72 characters in length. This includes white spaces.	
3 _____ Misaligned Amino Numbering	The numbering under each 5 th amino acid is misaligned. Do not use tab codes between numbers; use space characters, instead.	
4 _____ Non-ASCII	The submitted file was not saved in ASCII(DOS) text, as required by the Sequence Rules. Please ensure your subsequent submission is saved in ASCII text.	
5 _____ Variable Length	Sequence(s) _____ contain n's or Xaa's representing more than one residue. Per Sequence Rules, each n or Xaa can only represent a single residue. Please present the maximum number of each residue having variable length and indicate in the <220>-<223> section that some may be missing.	
6 _____ PatentIn 2.0 "bug"	A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequences(s) _____. Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. This applies to the mandatory <220>-<223> sections for Artificial or Unknown sequences.	
7 _____ Skipped Sequences (OLD RULES)	Sequence(s) _____ missing. If intentional, please insert the following lines for each skipped sequence: (2) INFORMATION FOR SEQ ID NO:X: (insert SEQ ID NO where "X" is shown) (i) SEQUENCE CHARACTERISTICS: (Do not insert any subheadings under this heading) (xi) SEQUENCE DESCRIPTION:SEQ ID NO:X: (insert SEQ ID NO where "X" is shown) This sequence is intentionally skipped Please also adjust the "(ii) NUMBER OF SEQUENCES:" response to include the skipped sequences.	
8 _____ Skipped Sequences (NEW RULES)	Sequence(s) _____ missing. If intentional, please insert the following lines for each skipped sequence. <210> sequence id number <400> sequence id number 000	
9 _____ Use of n's or Xaa's (NEW RULES)	Use of n's and/or Xaa's have been detected in the Sequence Listing. Per 1.823 of Sequence Rules, use of <220>-<223> is MANDATORY if n's or Xaa's are present. In <220> to <223> section, please explain location of n or Xaa; and which residue n or Xaa represents.	
10 _____ Invalid <213> Response	Per 1.823 of Sequence Rules, the only valid <213> responses are: Unknown, Artificial Sequence, or scientific name (Genus/species). <220>-<223> section is required when <213> response is Unknown or is Artificial Sequence	
11 _____ Use of <220>	Sequence(s) <u>10</u> missing the <220> "Feature" and associated numeric identifiers and responses. Use of <220> to <223> is MANDATORY if <213> "Organism" response is "Artificial Sequence" or "Unknown." Please explain source of genetic material in <220> to <223> section. (See "Federal Register," 06/01/1998, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of Sequence Rules)	
12 _____ PatentIn 2.0 "bug"	Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk.	
13 _____ Misuse of n	n can only be used to represent a single nucleotide in a nucleic acid sequence. N is not used to represent any value not specifically a nucleotide.	



RAW SEQUENCE LISTING

DATE: 01/24/2002

PATENT APPLICATION: US/09/586,744A

TIME: 10:19:50

Input Set : A:\9584-017.txt

Output Set: N:\CRF3\01242002\I586744A.raw

Does Not Comply
Corrected Diskette Needed

pp 1-4, 6

EX->

3 <110> APPLICANT: Harrington, et al.
 5 <120> TITLE OF INVENTION: Mammalian Flap Specific-Endonuclease
 7 <130> FILE REFERENCE: 9584-017
 9 <140> CURRENT APPLICATION NUMBER: 09/586,744A
 10 <141> CURRENT FILING DATE: 2000-06-02
 12 <160> NUMBER OF SEQ ID NOS: 74
 14 <170> SOFTWARE: PatentIn version 3.0
 16 <210> SEQ ID NO: 1
 17 <211> LENGTH: 380
 18 <212> TYPE: PRT
 19 <213> ORGANISM: Artificial
 21 <220> FEATURE:
 22 <223> OTHER INFORMATION: Peptide *give source of genetic material - see*
 24 <400> SEQUENCE: 1 *item 11 on*
 26 Met Gly Ile Gln Gly Leu Ala Lys Leu Ile Ala Asp Val Ala Pro Ser
 27 1 5 10 15
 29 Ala Ile Arg Glu Asn Asp Ile Lys Ser Tyr Phe Gly Arg Lys Val Ala
 30 20 25 30
 32 Ile Asp Ala Ser Met Ser Ile Tyr Gln Phe Leu Ile Ala Val Arg Gln
 33 35 40 45
 35 Gly Gly Asp Val Leu Gln Asn Glu Glu Gly Glu Thr Thr Ser His Leu
 36 50 55 60
 38 Met Gly Met Phe Tyr Arg Thr Ile Arg Met Met Glu Asn Gly Ile Lys
 39 65 70 75 80
 41 Pro Val Tyr Val Phe Asp Gly Lys Pro Pro Gln Leu Lys Ser Gly Glu
 42 85 90 95
 44 Leu Ala Lys Arg Ser Glu Arg Arg Ala Glu Ala Glu Lys Gln Leu Gln
 45 100 105 110
 47 Gln Ala Gln Ala Ala Gly Ala Glu Gly Glu Val Glu Lys Phe Thr Lys
 48 115 120 125
 50 Arg Leu Val Lys Val Thr Lys Gln His Asn Asp Glu Cys Lys His Leu
 51 130 135 140
 53 Leu Ser Leu Met Gly Ile Pro Tyr Leu Asp Ala Pro Ser Glu Ala Glu
 54 145 150 155 160
 56 Ala Ser Cys Ala Ala Leu Val Lys Ala Gly Lys Val Tyr Ala Ala Ala
 57 165 170 175
 59 Thr Glu Asp Met Asp Cys Leu Thr Phe Gly Ser Pro Val Leu Met Arg
 60 180 185 190
 62 His Leu Thr Ala Ser Glu Ala Lys Lys Leu Pro Ile Gln Glu Phe His
 63 195 200 205
 65 Leu Ser Arg Ile Leu Gln Glu Leu Gly Leu Asn Gln Glu Gln Phe Val
 66 210 215 220
 68 Asp Leu Cys Ile Leu Leu Gly Ser Asp Tyr Cys Glu Ser Ile Arg Gly

Encl
Summary sheet

RAW SEQUENCE LISTING

DATE: 01/24/2002

PATENT APPLICATION: US/09/586,744A

TIME: 10:19:50

Input Set : A:\9584-017.txt

Output Set: N:\CRF3\01242002\I586744A.raw

```

69 225          230          235          240
71 Ile Gly Pro Lys Arg Ala Val Asp Leu Ile Gln Lys His Lys Ser Ile
72          245          250          255
74 Glu Glu Ile Val Arg Arg Leu Asp Pro Asn Lys Tyr Pro Val Pro Glu
75          260          265          270
77 Asn Trp Leu His Lys Glu Ala His Gln Leu Phe Leu Glu Pro Glu Val
78          275          280          285
80 Leu Asp Pro Glu Ser Val Glu Leu Lys Trp Ser Glu Pro Asn Glu Glu
81          290          295          300
83 Glu Leu Ile Lys Phe Met Cys Gly Glu Lys Gln Phe Ser Glu Glu Arg
84 305          310          315          320
86 Ile Arg Ser Gly Val Lys Arg Leu Ser Lys Ser Arg Gln Gly Ser Thr
87          325          330          335
89 Gln Gly Arg Leu Asp Asp Phe Phe Lys Val Thr Gly Ser Leu Ser Ser
90          340          345          350
92 Ala Lys Arg Lys Glu Pro Glu Pro Lys Gly Ser Thr Lys Lys Lys Ala
93          355          360          365
95 Lys Thr Gly Ala Ala Gly Lys Phe Lys Arg Gly Lys
96          370          375          380

```

98 <210> SEQ ID NO: 2

99 <211> LENGTH: 1144

100 <212> TYPE: DNA

101 <213> ORGANISM: Artificial

103 <220> FEATURE:

104 <223> OTHER INFORMATION: CDNA *gene source*

106 <400> SEQUENCE: 2

```

107 atgggaattc aaggcctggc caaactaatt gctgatgtgg cccccagtgc catccgggag      60
109 aatgacatca agagctactt tggccgtaag gtggccattg atgcctctat gaggatttat      120
111 cagttcctga ttgctgttcg ccagggtggg gatgtgctgc agaatgagga gggtagagacc      180
113 accagccacc tgatgggcat gttctaccgc accattcgca tgatggagaa cggcatcaag      240
115 cccgtgtatg tctttgatgg caagccgcca cagctcaagt caggcgagct ggccaaacgc      300
117 agtgagcggc gggctgaggc agagaagcag ctgcagcagg ctcaggctgc tggggccgag      360
119 caggaggtgg aaaaattcac taagcggctg gtgaaggtca ctaagcagca caatgatgag      420
121 tgcaaacatc tgctgagcct catgggcatc ctttatcttg atgcacccag tgaggcagag      480
123 gccagctgtg ctgccctggt gaaggctggc aaagtctatg ctgcggctac cgaggacatg      540
125 gactgcctca ccttcggcag ccctgtgcta atgcgacacc tgactgccag tgaagccaaa      600
127 aagctgccaa tccaggaatt ccacctgagc cggattctgc aggagctggg cctgaaccag      660
129 gaacagtttg tggatctgtg catcctgcta ggcagtgact actgtgagag tatccggggt      720
131 attgggcccc agcgggctgt ggacctcatc cagaagcaca agagcatcga ggagatcgtg      780
133 cggcgacttg accccaacaa gtacctgtg ccagaaaatt ggctccacaa ggaggctcac      840
135 cagctcttct tggaaacctga ggtgctggac ccagagtctg tggagctgaa gtggagcgag      900
137 ccaaatgaag aagagctgat caagttcatg tgtggtgaaa agcagttctc tgaggagcga      960
139 atccgcagtg gggtaagag gctgagtaag agccgccaa ggcagaccca gggccgctg      1020
141 gatgatttct tcaaggtgac cggtcactc tcttcagcta agcgcaagga gccagaaccc      1080
143 aagggatcca ctaagaagaa ggcaaagact ggggcagcag ggaagtttaa aaggggaaaa      1140
145 taaa

```

148 <210> SEQ ID NO: 3

149 <211> LENGTH: 377

150 <212> TYPE: PRT

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/586,744A

DATE: 01/24/2002

TIME: 10:19:50

Input Set : A:\9584-017.txt

Output Set : N:\CRF3\01242002\I586744A.raw

151 <213> ORGANISM: Artificial
 153 <220> FEATURE:
 154 <223> OTHER INFORMATION: Peptide *givi source*
 156 <400> SEQUENCE: 3
 158 Met Gly Ile His Gly Leu Ala Lys Leu Ile Ala Asp Val Ala Pro Ser
 159 1 5 10 15
 161 Ala Ile Arg Glu Asn Asp Ile Lys Ser Tyr Phe Gly Arg Lys Val Ala
 162 20 25 30
 164 Ile Asp Ala Ser Met Ser Ile Tyr Gln Phe Leu Ile Ala Val Arg Gln
 165 35 40 45
 167 Gly Gly Asp Val Leu Gln Asn Glu Glu Gly Glu Thr Thr Ser Leu Met
 168 50 55 60
 170 Gly Met Phe Tyr Arg Thr Ile Arg Met Glu Asn Gly Ile Lys Pro Val
 171 65 70 75 80
 173 Tyr Val Phe Asp Gly Lys Pro Pro Gln Leu Lys Ser Gly Glu Leu Ala
 174 85 90 95
 176 Lys Arg Ser Glu Arg Arg Ala Glu Ala Glu Lys Gln Leu Gln Gln Ala
 177 100 105 110
 179 Gln Glu Ala Gly Met Glu Glu Val Glu Lys Phe Thr Lys Arg Leu Val
 180 115 120 125
 182 Lys Val Thr Lys Lys Gln His Asn Asp Glu Cys Lys His Leu Leu Ser Leu
 183 130 135 140
 185 Met Gly Ile Pro Tyr Leu Asp Ala Pro Ser Glu Ala Glu Ala Ser Cys
 186 145 150 155 160
 188 Ala Ala Leu Ala Lys Ala Gly Lys Val Tyr Ala Ala Ala Thr Glu Asp
 189 165 170 175
 191 Met Asp Cys Leu Thr Phe Gly Ser Pro Val Leu Met Arg His Leu Thr
 192 180 185 190
 194 Ala Ser Glu Ala Lys Lys Leu Pro Ile Gln Glu Phe His Leu Ser Arg
 195 195 200 205
 197 Val Leu Gln Glu Leu Gly Leu Asn Gln Glu Gln Phe Val Asp Leu Cys
 198 210 215 220
 200 Ile Leu Leu Gly Ser Asp Tyr Cys Glu Ser Ile Arg Gly Ile Gly Ala
 201 225 230 235 240
 203 Lys Arg Ala Val Asp Leu Ile Gln Lys His Lys Ser Ile Glu Glu Ile
 204 245 250 255
 206 Val Arg Arg Leu Asp Pro Ser Lys Tyr Pro Val Pro Glu Asn Trp Leu
 207 260 265 270
 209 His Lys Glu Ala Gln Gln Leu Phe Leu Glu Pro Glu Val Val Asp Pro
 210 275 280 285
 212 Glu Ser Val Glu Leu Lys Trp Ser Glu Pro Asn Glu Glu Glu Leu Val
 213 290 295 300
 215 Lys Phe Met Cys Gly Glu Lys Gln Phe Ser Glu Glu Arg Ile Arg Ser
 216 305 310 315 320
 218 Gly Val Lys Arg Leu Ser Lys Ser Arg Gln Gly Ser Thr Gln Gly Arg
 219 325 330 335
 221 Leu Asp Asp Phe Phe Lys Val Thr Gly Ser Leu Ser Ser Ala Lys Arg
 222 340 345 350
 224 Lys Glu Pro Glu Pro Lys Gly Ser Ala Lys Lys Lys Ala Lys Thr Gly

RAW SEQUENCE LISTING

DATE: 01/24/2002

PATENT APPLICATION: US/09/586,744A

TIME: 10:19:50

Input Set : A:\9584-017.txt

Output Set: N:\CRF3\01242002\I586744A.raw

225 355 360 365
 227 Gly Ala Gly Lys Phe Arg Arg Gly Lys
 228 370 375
 230 <210> SEQ ID NO: 4
 231 <211> LENGTH: 1930
 232 <212> TYPE: DNA
 233 <213> ORGANISM: Artificial
 235 <220> FEATURE:
 236 <223> OTHER INFORMATION: CDNA
 238 <400> SEQUENCE: 4
 239 atgggaattc acggccttgc caaactaatt gctgatgtgg cccccagtgc catccgtgag 60
 241 aatgacatca agagctactt tggtcgtaaa gtggccatcg atgcctccat gagcatctac 120
 243 cagttcctga ttgctgttcg tcagggtggg gatgtgtcgc agaacgagga ggggtgagacc 180
 245 accagcctga tgggcatgtt atggcaaacc atccgcatgg agaatggcat caagcctgtg 240
 247 tacgtctttg atggcaaacc accacagctg aagtcaggcg agctggccaa gcgcagtgag 300
 249 aggcgcgcgc aggctgagaa gcaactgcag caggctcagg aggctgggat ggaggaggag 360
 251 gtggagaagt tcaccaagag gctcgtgaag gtcaccaagc aacacaatga tgagtgc aaa 420
 253 caoctcgtga gcctcatggg catcccttac cttgatgcac ccagecaggc agaggccagc 480
 255 tgtgctgccc tggcaaaggc tggcaaagtc tatgctgcgg ccacggagga catggactgc 540
 257 ctcaactttt gcagccccgt gctaattgca cacttaactg ccagtgaggc caagaagctg 600
 259 cccatccaag agttccatct gagccgcgtc ctgcaggagc tgggtctgaa ccaggagcag 660
 261 tttgtggatc tgtgcatect gctgggtagc gactactgcg agagcatccg tggcattggc 720
 263 gccaaagcgg ctgtggatct catccagaaa cataagagca tcgaggagat cgtgaggcgg 780
 265 ctggacccca gcaagtaccc cgttccagag aactggctcc acaagggaagc ccagcagctc 840
 267 ttctctggagc cagaagtagt ggacccagag tctgtggagc tgaagtggag cgagccaaat 900
 269 gaagaagagt tgggtcaaatt tatgtgtggt gaaaagcagt tttctgaaga gcgaattcgc 960
 271 agtgggggtc agcggctgag taagagccgc cagggcagca cccagggagc cctcgatgat 1020
 273 ttcttcaagg tgacaggctc actctcctca gctaagcgca aggagccaga acccaagggg 1080
 275 cctgctaaga agaaagcaaa gactggggga gcgggggaagt tccgaagggg aaaataaacc 1140
 277 tgtccttccc ctccactgtc cttgacccca ggctgtctat ctgttttgta ccctgcgctg 1200
 279 cagcacatcc ctcttgtccc tegtcttgag gagagttcat tgcttccagc gctcgccttc 1260
 281 agagctttcc ctctcttgac cctgtggcag gaaggccgta gctctgcttt ttctcatttt 1320
 283 tagctcagga aagatgtcag gctcaaacca cttctcaggt taatggacac tgtagtcatt 1380
 285 gttctgtgca actgcgagca atgtcttaag gaagaagaag ataaagccgg gagcgaggct 1440
 287 ggagatagtt tcccagctgg ccagctggtg gaggagaggt gactagaacc tgactgacta 1500
 289 ctgctccttc taatttcaact gtccctgaaa gatgccatc agcctgggat tcgctgatgg 1560
 291 aagaactgca aagagacgca gcagagagaa gtctggctga caacagattt agtactgacc 1620
 293 agctgatttt tgtgggcaga aatttgaact tgctgcctgc tgagtccagt agttgtgcag 1680
 295 ggagtgaagt ggcagtgatt aagttttgat ttgtagtttt ttgtttttgt ctctccctc 1740
 297 tccagtgatt gggatggacc ccagggcaaa ggcattaagt gtgccactga cctgtgcctc 1800
 299 caagtgaatg tctgacagcc tttctgaggc aatcaattga attgaggttt tgggagaaga 1860
 301 aactgttggt cataggctat ttctatttta aaagatgtga agagaaaaaa aaaacaataa 1920
 303 aattataaaa
 306 <210> SEQ ID NO: 5
 307 <211> LENGTH: 382
 308 <212> TYPE: PRT
 309 <213> ORGANISM: Artificial
 311 <220> FEATURE:
 312 <223> OTHER INFORMATION: Peptide

Please correct this error in subsequent sequences.

RAW SEQUENCE LISTING

DATE: 01/24/2002

PATENT APPLICATION: US/09/586,744A

TIME: 10:19:50

Input Set : A:\9584-017.txt

Output Set: N:\CRF3\01242002\I586744A.raw

314 <400> SEQUENCE: 5

```

316 Met Gly Ile Lys Gly Leu Asn Ala Ile Ile Ser Glu His Val Pro Ser
317 1      5      10      15
319 Ala Ile Arg Lys Ser Asp Ile Lys Ser Phe Phe Gly Arg Lys Val Ala
320      20      25      30
322 Ile Asp Ala Ser Met Ser Leu Tyr Gln Phe Leu Ile Ala Val Arg Gln
323      35      40      45
325 Gln Asp Gly Gly Gln Leu Thr Asn Glu Ala Gly Glu Thr Thr Ser His
326      50      55      60
328 Leu Met Gly Met Phe Tyr Arg Thr Leu Arg Met Ile Asp Asn Gly Ile
329 65      70      75      80
331 Lys Pro Cys Tyr Val Phe Asp Gly Lys Pro Pro Asp Leu Lys Ser His
332      85      90      95
334 Glu Leu Thr Lys Arg Ser Ser Arg Arg Val Glu Thr Glu Lys Lys Leu
335      100     105     110
337 Ala Glu Ala Thr Thr Glu Leu Glu Lys Met Lys Gln Glu Arg Arg Leu
338      115     120     125
340 Val Lys Val Ser Lys Glu His Asn Glu Glu Ala Gln Lys Leu Leu Gly
341      130     135     140
343 Leu Met Gly Ile Pro Tyr Ile Ile Ala Pro Thr Glu Ala Glu Ala Gln
344 145     150     155     160
346 Cys Ala Glu Leu Ala Lys Lys Gly Lys Val Tyr Ala Ala Ala Ser Glu
347      165     170     175
349 Asp Met Asp Thr Leu Cys Tyr Arg Thr Pro Phe Leu Leu Arg His Leu
350      180     185     190
352 Thr Phe Ser Glu Ala Lys Lys Glu Pro Ile His Glu Ile Asp Thr Glu
353      195     200     205
355 Leu Val Leu Arg Gly Leu Asp Leu Thr Ile Glu Gln Phe Val Asp Leu
356      210     215     220
358 Cys Ile Met Leu Gly Cys Asp Tyr Cys Glu Ser Ile Arg Gly Val Gly
359 225     230     235     240
361 Pro Val Thr Ala Leu Lys Leu Ile Lys Thr His Gly Ser Ile Glu Lys
362      245     250     255
364 Ile Val Glu Phe Ile Glu Ser Gly Glu Ser Asn Asn Thr Lys Trp Lys
365      260     265     270
367 Ile Pro Glu Asp Trp Pro Tyr Lys Gln Ala Arg Met Leu Phe Leu Asp
368      275     280     285
370 Pro Glu Val Ile Asp Gly Asn Glu Ile Asn Leu Lys Trp Ser Pro Pro
371      290     295     300
373 Lys Glu Lys Glu Leu Ile Glu Tyr Leu Cys Asp Asp Lys Lys Phe Ser
374 305     310     315     320
376 Glu Glu Arg Val Lys Ser Gly Ile Ser Arg Leu Lys Lys Gly Leu Lys
377      325     330     335
379 Ser Gly Ile Gln Gly Arg Leu Asp Gly Phe Phe Gln Val Val Pro Lys
380      340     345     350
382 Thr Lys Glu Gln Leu Ala Ala Ala Ala Lys Arg Ala Gln Glu Asn Lys
383      355     360     365
385 Lys Leu Asn Lys Asn Lys Asn Lys Val Thr Lys Gly Arg Arg
386      370     375     380

```

09/586,744A 6

<210> 10

<211> 378

<212> PRT

<213> Artificial Sequence

<220> !

<223>

<400> 10

→ see item 11 on Ena Summary sheet

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/586,744A

DATE: 01/24/2002

TIME: 10:19:51

Input Set : A:\9584-017.txt

Output Set: N:\CRF3\01242002\I586744A.raw

L:19 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:1
L:101 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:2
L:151 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:3
L:233 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:4
L:309 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:5
L:391 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:6
L:441 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:7
L:526 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:8
L:576 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:9
L:718 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:10
L:720 M:258 W: Mandatory Feature missing, <220> FEATURE:
L:720 M:258 W: Mandatory Feature missing, <223> OTHER INFORMATION:
L:821 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:11
L:833 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:12
L:846 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:13
L:858 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:14
L:870 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:15
L:886 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:16
L:899 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:17
L:915 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:18
L:931 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:19
L:947 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:20
L:963 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:21
L:979 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:22
L:995 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:23
L:1011 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:24
L:1027 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:25
L:1043 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:26
L:1059 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:27
L:1072 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:28
L:1122 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:29
L:1134 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:30
L:1146 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:31
L:1158 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:32
L:1170 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:33
L:1182 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:34
L:1194 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:35
L:1206 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:36
L:1218 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:37
L:1230 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:38
L:1242 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:39
L:1254 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:40
L:1266 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:41
L:1278 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:42
L:1290 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:43
L:1302 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:44
L:1314 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:45
L:1326 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:46

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/586,744A

DATE: 01/24/2002

TIME: 10:19:51

Input Set : A:\9584-017.txt

Output Set: N:\CRF3\01242002\I586744A.raw

L:1338 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:47
L:1350 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:48
L:1362 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:49
L:1374 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:50